An Integrated Apicomplexan Parasite Database

Jessica Kissinger (Univ. Georgia)
David Roos (Univ. Pennsylvania)
Chris Stoeckert (Univ. Pennsylvania)

PlasmoDB
The Plasmodium Genome Resource
Release 4.0
Data Release: 10/03/2002

PlasmoDB is the official database of the malaria parasite genome project. This resource provides access to finished sequence for *Plasmodium falciparum* (strain 3D7), generated by the Sanger Institute, TIGR/NMRC, and Stanford, in addition to a wide range of related genomic-scale datasets from other sources. Publications exploiting PlasmoDB should provide appropriate acknowledgement to the database developers and those scientists who have made their data available on this site (see Acknowledgements and Publications). Development of PlasmoDB has been supported by the Burroughs Wellcome Fund.

PlasmoDB Downloads Tools Queries History CDs&Links Browse Sources SRT Help

Latest News:
- Finished sequence and annotation released.
- PlasmoDB 4.0 released.
- More News ...

Quick Links:
- Browser Settings
- Tutorials
- Publications
- BLAST
- Sequence Retrieval Tool (SRT)
- Download Sequences

P. falciparum go!

Features:
- Queries:

Welcome to PlasmoD

PlasmoDB 4.0 provides a greatly enhanced with publication of the complete parasite ge PlasmoDB 4.0 incorporates DNA sequence sequencing centers (Nature 419:498-511); au algorithms; cDNA sequences (mapped to th motifs, T-cell epitopes, etc; cross-species co optical and genetic mapping data; population generated by a variety of complementary str information relevant to malaria research. A v manipulation, ePCR, BLAST searching, mot mass fingerprinting, etc. PlasmoDB is based a rich schema, enabling sophisticated queries

The blue navigation bar provided at the top components of PlasmoDB. The forms provi access some of the most commonly used pag with questions, problems, and suggestions fc Click here to see a list of features that we ex

Find an Annotated Gene by W

CryptoDB
The Cryptosporidium Genome Resources
Release 1.0
Data Release: 2003/01/14

CryptoDB provides access to the draft sequence for the genome of *C. parvum* (type I and II), generated by VCU and MINNESOTA, as well as additional *C. parvum* data from various other sources (see Acknowledgements). Data are available for browsing, querying, and datamining subject to the Data Release Policy.

ome to CryptoDB 1.0
ncludes unfinished genome sequence data. Analysis of poses many difficulties. In particular, some of the data has not remove vector, bacterial, host or other contaminating sequences!

ls ▸
- .AST
- d a Protein Motif

ries ▸
- xt Search

wse the Data ▸
- ta Sources
- edline Cryptosporidium
- dates
- yptosporidium Links

Help ▸
- Tutorial on CryptoDB
- Sequence Naming Convention
- Frequently Asked Questions
- What's New *new*
- Known Bugs
- Methods
- Acknowledgements
- How to cite CryptoDB and Data Sources

ToxoDB
The Toxoplasma Genome Resource
Release 2.2
Data Release: 2003/05/26

ToxoDB provides access to the draft sequence for the genome of *T. gondii* (ME49, a type II strain), generated by TIGR, as well as additional *T. gondii* data from various other sources (see Acknowledgements). Data are available for browsing, querying, and datamining subject to the Data Release Policy.

Welcome to ToxoDB 2.2

WARNING: This database includes unfinished genome sequence data. Analysis of unfinished data poses many difficulties. In particular, some of the data has not been filtered to remove vector, bacterial, host or other contaminating sequences!

WARNING: As of 05/26/03, 14:30 EST the new release of ToxoDB is up (after last-minute work done from Tarrytown, NY!). However, please note that this site is up without much testing. So, please be patient, send email to report bugs, and watch for corrections. Thanks!

Quick Links ▸

Sequence Retrieval Tool (SRT) GO!

Download Bulk Sequence *new data* GO!

Download Single Sequence:
sequence ID TGG_1095
Start 1 End 200 GO!

Organellar Genome: Plastid GO!

Tools ▸
- BLAST
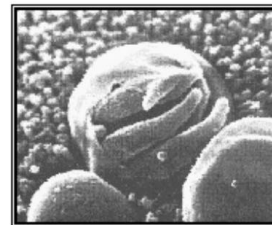- Find a Protein Motif
- Gene2EST
- Lutefisk
- EMOWSE

Queries ▸
- Text Search

Browse the Data ▸
- Data Sources
- ApiESTDB *new*
- Medline Toxo Updates
- Toxo Links

Help ▸
- Tutorial on ToxoDB
- Sequence Naming Convention
- Frequently Asked Questions
- What's New
- Known Bugs
- Methods
- Acknowledgements
- How to cite ToxoDB and Data Sources

adapted from *Sogin et al* (1991)

Animals

Fungi

Plants

superphylum Alveolata (kingdom Protozoa)

10 changes

100 — **Plasmodium vivax**
75 — *P. knowlesi*
85 — *P. berghei*
— *P. malariae*
100
95 — **P. falciparum**
— *P. reichenowi*
48 — *P. gallinaceum*

100 — *Babesia equi*
— *Theileria bufeii*

100 — *Sarcocystis gigante*
— **Toxoplasma gondii**

Cryptosporidium *wrairi*

Apicomplexa
5000 known species of protozoan parasites

100 — *Prorocentrum micans*
— *Symbiodinim microadriaticum*

Dinoflagellata

100 — *Paramecium tetraurelia*
— *Oxytricha nova*

Ciliophora

Distinctive Features of (most) Apicomplexan Parasites

- *Complex developmental life cycle involving temporal differentiation, latent cysts*

- *Intracellular pathogen; parasitophorous vacuole mediates interaction with the host cell*

- *Replicates by assembling daughter parasites within the mother (schizogony)*

- *Standard eukaryotic machin-ery, plus unusual organelles:*

- *rhoptries & micronemes (for host cell attachment/invasion)*

- *conoid (apical cytoskeleton)*

- *inner membrane complex (assembly of daughter cells)*

- *apicoplast (2° endosymbiont)*

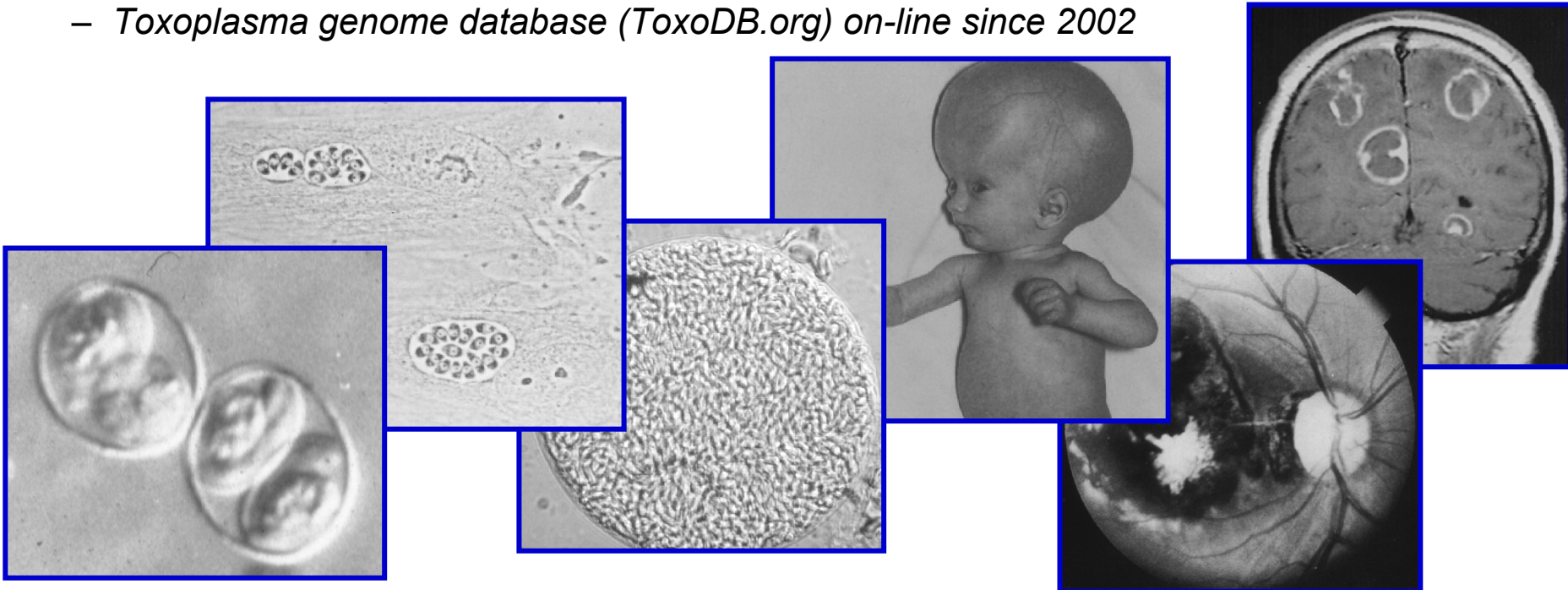# *Cryptosporidium* species – class B biodefense agent

- **Biology:** *common intestinal pathogen*
- **Clinical:** *immunosuppressed individuals, no treatment*
- **Research:** *laboratory study extremely difficult*
- **Genomics (and beyond)**
  - *9 Mb C. parvum sequence completed;*
    C. hominis sequence effectively complete
  - <u>*highly*</u> *reduced metabolic pathways;*
    many horizontally-transferred genes
    including potential drug targets);
    multiple transporters
  - *few ESTs (~1K); arrays in planning?;*
    proteomics difficult
  - *Cryptosporidium genome database*
    (CryptoDB.org) on-line since 2004

Toxoplasma gondii – class B biodefense agent

- **Biology:** *ubiquitous human pathogen*
- **Clinical:** *congenital infection, immunosuppressed individuals*
- **Research:** *accessible model organism*
- **Genomics (and beyond)**
  - *~65 Mb reference sequence effectively complete, seq. of two additional strains proposed*
  - *~100K ESTs, ~400K SAGE tags, representing all major strains, 3 life-cycle stages*
  - *cDNA-based glass slide arrays in use, photolithographic arrays under development*
  - *some proteomics data available, especially for apicomplexan-specific organelles*
  - *Toxoplasma genome database (ToxoDB.org) on-line since 2002*

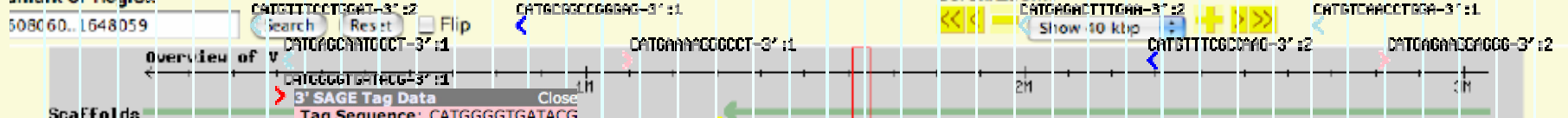# ToxoDB

Showing 40 kbp from V, positions 1,608,060 to 1,648,059

Instructions: Search using a sequence name, gene name, locus, or other landmark. The wildcard character * is allowed. To center on a location, click the ruler. Use the Scroll/Zoom buttons to change magnification and position.
Examples: dynein heavy chain, Chromosome:IV, Scaffold:995283, Contig:994378, GeneticMarker:SAG4, tRNA:Lys*.

[Hide banner] [Hide instructions] [Bookmark this view] [Link to an image of this view] [Publication quality image] [Help]

Landmark or Region

V:1608060..1648059   (Search) (Reset) ☐ Flip

Scroll/Zoom:
<< < — Show 40 kbp ⊡ + > >>

Overview of V

1M   2M   3M

Scaffolds

Markers

KT-850   AK24   KT-354   L358   KT-cA5 pS1-13 AK25
     AK22   AK14

Data Source
T. gondii Genome Data ⊡

Dumps, Searches and other Operations:
Annotate Restriction Sites ⊡ (About...) (Configure...) (Go)

Tracks [Hide]

(External tracks italicized)
*Overview track

☑ Scaffolds*	☐ 3' Redundant SAGE Tags	☐ BLASTX Archaea
☑ Markers*	☐ 5' Redundant SAGE Tags	☐ *T. gondii UniGene EST Alignments*
☐ Markers	☐ tRNAscan-SE Predictions	☐ *TIGR T. gondii Gene Index*
☑ Scaffolds	☑ TwinScan Gene Predictions	☑ *T. gondii dbEST Alignments*
☐ Contigs	☑ GlimmerHMM Gene Predictions	☐ *Coccidia dbEST Alignments*
☑ Sanger BAC Ends	☑ TigrScan Gene Predictions	☐ *Apicomplexa dbEST Alignments*
☐ Repeat Regions	☑ GeneFinder Gene Predictions	☐ 3-frame translation (forward)
☐ Annotated Genes	☐ BLASTX T. gondii	☐ DNA/GC Content
☐ Annotation PCR Primers	☑ BLASTX Apicomplexa	☐ 3-frame translation (reverse)
☑ 3' SAGE Tags	☑ BLASTX Eukaryota	☐ plugin:Restriction Sites
☐ 5' SAGE Tags	☐ BLASTX Bacteria	

Image Width
☐ 450 ☐ 640 ☐ 800 ☑ 1024

Key position
☑ Between ☐ Beneath ☐ Left ☐ Right

Track Name Table
☐ Alphabetic ☑ Varying

(Set Track Options...) (Update Image)

Upload your own annotations: [Help]
Upload a file (Choose File) no file selected   (Upload) (New...)

Add remote annotations: [Help]
Enter Remote Annotation URL [   ]
(Update URLs)

For the source code for this browser, see the Generic Model Organism Database Project. For other questions, send mail to help@toxodb.org.

ToxoDB

Click here to go to the ToxoDB homepage.

Showing 40 kbp from V, positions 1,608,060 to 1,648,059

Instructions: Search using a sequence name, gene name, locus, or other landmark. The wildcard character * is allowed. To center on a location, click the ruler. Use the Scroll/Zoom buttons to change magnification and position.
Example: dynein heavy chain, Chromosome:IV, Scaffold:99528*, Contig 994179, GeneticMarker:SAG4, tRNA:Lys*.

[Hide banner] [Hide instructions] [Bookmark this view] [Link to an image of this view] [Publication quality image] [Help]

Landmark or Region:
1608060..1648059 | Search | Reset | ☐ Flip

Scroll/Zoom: |<< |< — < Show 40 kbp ⇕ > — > >>|

Scaffolds

3' SAGE Tag Data Close
Tag Sequence: CATGGGGTGATACG
# of Matches: 1
SAGE Library Raw Frequencies:

SP	D4	D6	D7	D17	RH	MSJ	B7
4	8	2	6	11	1	3	0

Lookup Tag in TgSAGE DB

TwinScan Gene Predictions
TgTwinScan_3568 TgTwinScan TgTwinScan_3564 TgTwinScan_3563 TgTwinScan_3562 TgTwinScan_35

GlimmerHMM Gene Predictions
TgGlmHMM_4930 TgGlmHMM_4928 TgGlmHMM_4927 TgGlmHMM_4926 TgGlmHMM_4924
TgGlmHMM_4929 TgGlmHMM_4925

TigrScan Gene Predictions
TgTigrScan_3126 TgTigrScan_3124 TgTigrScan_3123 TgTigrScan_3122 TgTigrScan_3121
TgTigrScan_3125 TgTigrScan_3120

GeneFinder Gene Predictions
TgGeneFinder_14075 TgGeneFinder_14074 TgGeneFinder_14073 TgGeneFinder_14066 TgGeneFinder_14072
TgGeneFinder_14067

BLASTX Apicomplexa

BLASTX Eukaryota

BLASTX Alignment Data
ProteinMatchID: 1247113
gi|23478454|gb|EAA15538.1|
hypothetical protein [Plasmodium yoeli
Show BlastX alignment
%ID: 29
E(): 9.6e-07
Length: 258
Span: 7-195

T. gondii dbEST Alignments

Plasmodium species – (re)emerging pathogens

- ● Biology: *extreme specialist*; *important force in human evolution*
- ● Clinical: *problem acute & growing; drug-resistance, no vaccine*
- ● Research: *laboratory study now possible (but difficult)*
- ● Genomics (and beyond)

  – *23 Mb genome P. falciparum sequence completed; complete sequencing of P. vivax in progress; 3-5X coverage available for several other species (P. berghei, P. chabaudi, P. gallinaceum, P. knowlesi, P. yoelii); sequencing of field isolate planned*

  – *24K ESTs available for P.falciparum, 1K for P.vivax, 25K for rodent malaria species (P. berghei & P. yoelii); some SAGE tag data, high quality expression profiling data from both glass-slide and photolithographic oligonucleotide arrays*

  – *Plasmodium genome database (PlasmoDB.org) on-line since 2000*

PlasmoDB
The Plasmodium Genome Resource

Release 4.1
Data Release:
01/23/2003
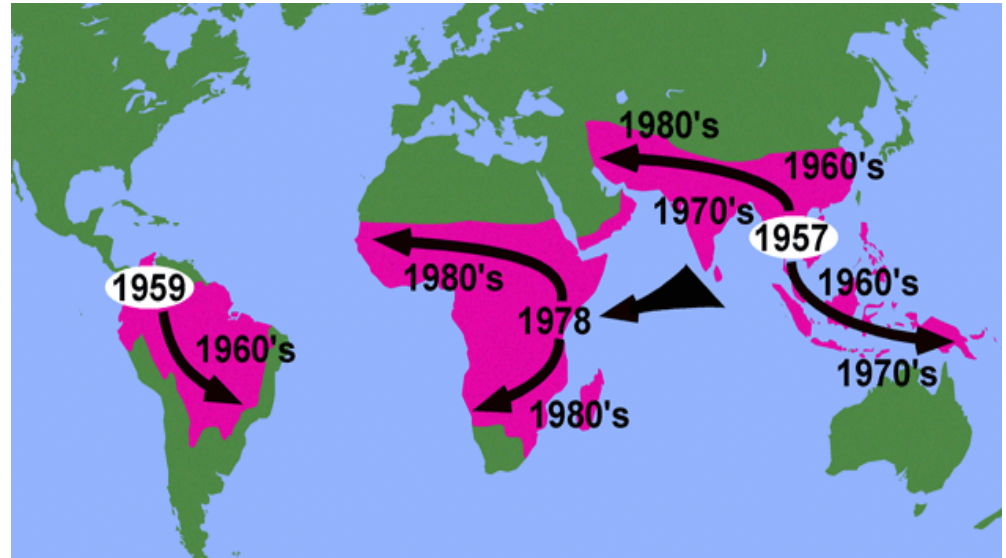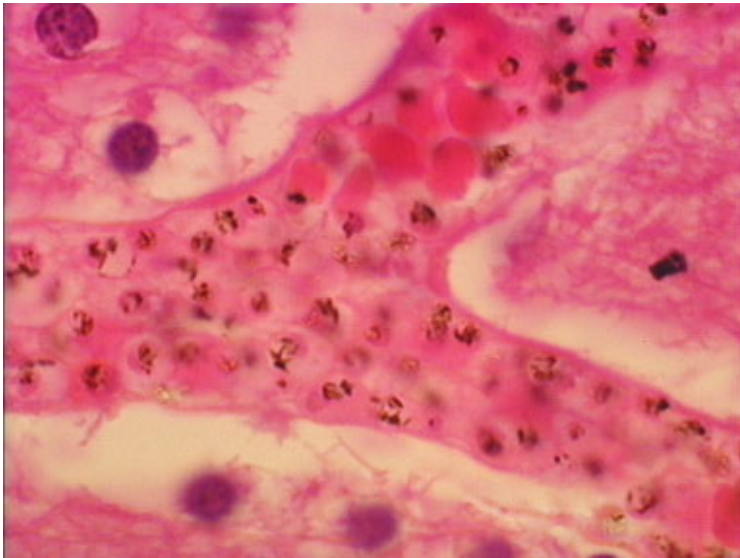
Genes whose protein contains a predicted signal

Query result: rows 1 - 100

| Home | Downloads | Tools | Queries | BLAST | History | CDs & Links | Browse | Data Sources | SRT |

Query: Genes [type=sequencing center annotations (Pf Annotation) chr=all] predicted to contain a signal peptide.

[1-100][**101-200**]   Rows 1 - 100 of 651   [1-100][10

	gene	location	description
1	PFA0025c	pfal_chr1: 53392-53503	VAR fragment, pseudogene
2	PFA0030c	pfal_chr1: 54001-55229	rifin
3	PFA0050c	pfal_chr1: 66050-67222	rifin
4	PFA0060w	pfal_chr1: 71857-72659	hypothetical protein, conserved in P. falciparum
5	PFA0090c	pfal_chr1: 87436-88410	stevor
6	PFA0095c	pfal_chr1: 90475-91653	rifin
7	PFA0125c	pfal_chr1: 110984-116033	Ebl-1 like protein, putative
8	PFA0135w	pfal_chr1: 124752-125719	hypothetical protein
9	PFA0180w	pfal_chr1: 161365-166464	hypothetical protein
10	PFA0195w	pfal_chr1: 173099-174826	hypothetical protein
11	PFA0210c	pfal_chr1: 183057-184457	hypothetical protein
12	PFA0225w	pfal_chr1: 202774-204381	LytB protein

PlasmoDB
The Plasmodium Genome Resource

Release 4.1
Data Release:
01/23/2003

Genes with a specified phylogenetic profile

Query result: rows 1 - 20

Home   Downloads   Tools   Queries   BLAST   History   CDs & Links   Browse   Data Sources   SRT

Query: All Pf genes with the following phylogenetic profile: *A. thaliana*=don't care *C. elegans*=don't care *D. melanogaster*=don't care *E. coli*=don't care *H. sapiens*=no *S. cerevisiae*=don't care *A. gambiae*=don't care *M. musculus*=don't care *P. yoelii*=yes

[1-20][21-40]   Rows 1 - 20 of 2260   [1-20][21

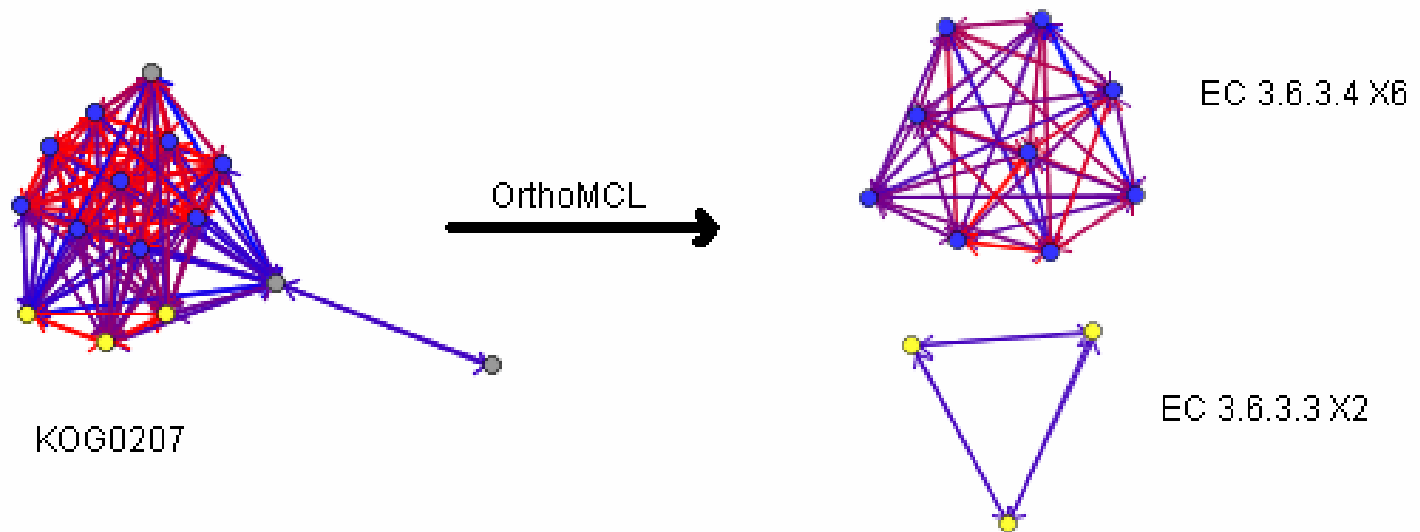|	gene	ortholog group	group size	location	description
1	PF11_0274	755634	2	chr11: 1027592-1028394	hypothetical protein
2	PF10_0083	755635	2	chr10: 357345-358463	hypothetical protein
3	PF11_0275	755637	2	chr11: 1029975-1034081	hypothetical protein
4	PF10_0151	755639	2	chr10: 621303-623624	hypothetical protein
5	MAL6P1.56	755641	2	chr6: 272734-274282	ST kinase, putative
6	PF10_0154	755642	2	chr10: 633681-635285	ribonucleotide reductase small subunit, putative
7	PFE0155w	755644	2	pfal_chr5: 130453-133392	hypothetical protein
8	PF11_0204	755645	2	chr11: 742488-743546	hypothetical protein
9	MAL6P1.298	755647	2	chr6: 576383-577498	hypothetical protein
10	PF11_0345	755648	2	chr11: 1293712-1295110	hypothetical protein

Home

Descri
Putativ
mosqu
can be
query
the pro
orthol
profile.

Query

Query

Identification of ortholog groups using OrthoMCL

Protein sequences from multiple organisms

All-against-all BLASTP

Within species:
reciprocal **better** similarity pairs define putative in-paralogs

Between species:
reciprocal best similarity pairs define putative orthologs

Similarity Matrix
Normalized edge weights

Markov Clustering

Ortholog / In-paralog groups

Li et al, Genome Research **13:***2178-90 (2003)*
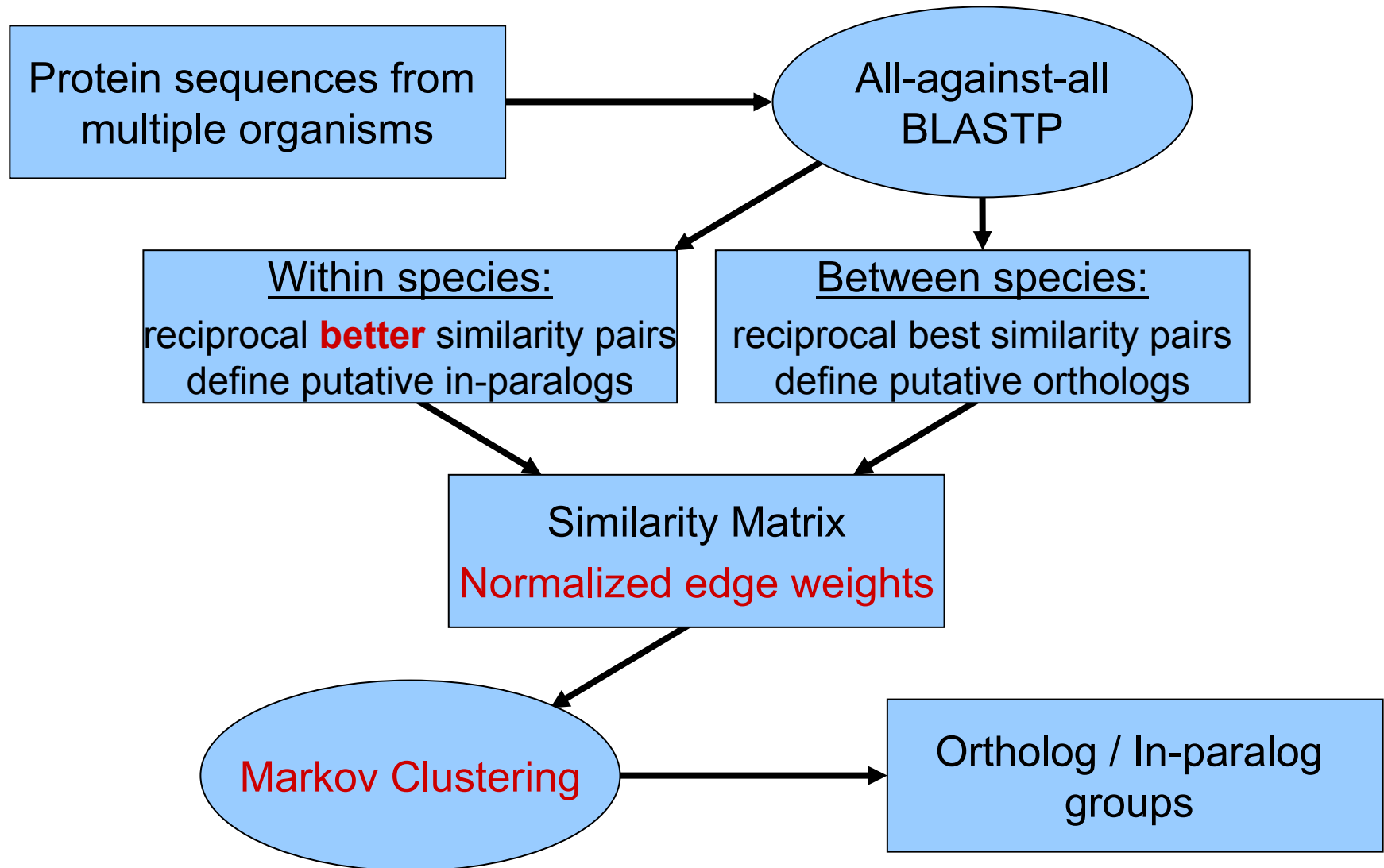
EC-consistency of OrthoMCL and KOG groups

Methods	Total dataset		Groups with ≥2 proteins for which complete EC annotations are available			Consistent EC assignments		
	groups	proteins (% of proteome)	groups	proteins	EC-annotated (% of total)	groups (% poss.)	proteins	EC-annotated (% possible)
OrthoMCL	**13253**	**78998 (69)**	**998**	**10460**	**4122 (86)**	**886 (88)**	**8123**	**3495 (84)**
KOG	**10058**	**88645 (78)**	**926**	**14471**	**4393 (92)**	**772 (83)**	**9927**	**3359 (76)**

KOG0207 → OrthoMCL → EC 3.6.3.4 X6 / EC 3.6.3.3 X2

Li Li & Feng Chen, unpublished

PlasmoDB
The Plasmodium Genome Resource

Release 4.1
Data Release: 01/23/2003

Scripps/GNF malaria array - genes ranked by expression

Query result: rows 1 - 20

| Home | Downloads | Tools | Queries | BLAST | History | CDs & Links | Browse | Data Sources | SRT |

Query: Genes [type=sequencing center annotations (Pf Annotation) chr=all] expressed in the 95th percentile or above in the Late Schizogony stage

[1-20][21-40]   Rows 1 - 20 of 247   [1-20][

	gene	avg. intensity	percentile	location	description
1	PFB0120w	15974.90	99.96	pfal_chr2: 127994-128314	hypothetical protein
2	PF11_0040	15559.15	99.94	chr11: 129320-129604	early transcribed membrane protein 11.2
3	PFB0300c	8217.25	99.92	pfal_chr2: 273689-274507	merozoite surface protein 2 precursor
4	PF10_0372	6192.70	99.90	chr10: 1508412-1509473	hypothetical protein
5	PF13_0058	5204.35	99.88	chr13_1: 426154-426585	hypothetical protein
6	PF14_0598	5187.70	99.86	chr14: 2558046-2559295	glyceraldehyde-3-phosphate dehydrogen
7	PF10_0019	5053.90	99.84	chr10: 81417-81740	early transcribed membrane protein
8	PFC0120w	4901.05	99.82	pfal_chr3: 132097-137339	Cytoadherence linked asexual protein, CL
9	PFA0420w	4622.80	99.80	pfal_chr1: 350323-350862	hypothetical protein
10	PFE0165w	4435.15	99.79	pfal_chr5: 140710-141471	actin depolymerizing factor, putative
11	MAL13P1.308	4349.45	99.77	chr13_1: 2369502-2377598	hypothetical protein
12	PF11_0039	4346.10	99.75	chr11: 126406-126681	early transcribed membrane protein 11.1
13	PF11_0224	4332.25	99.73	chr11: 813000-813936	circumsporozoite-related antigen

Available microarray datasets spanning the P. falciparum intraerythrocytic life cycle
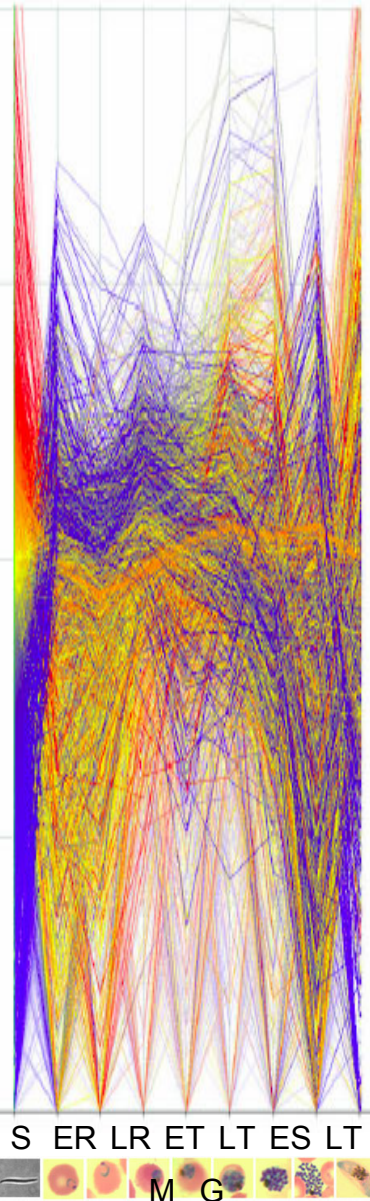
	Le Roch *et al*	Bozdech *et al*
Array Platform	Custom Affy	Glass Slide
Oligos Mapped to Pred Coding Seq	260,596	7462
Oligos Mapped to Pred Non-coding Seq	106,630	None
Average # of Oligos Used per Gene	17.4	1.3
Genes for which Data is Available	5104	4372
Erythrocytic Time Points	6 + 6	48 + 50 + 53
P. falciparum Strain	3D7	HB3/DD2/3D7
Synchronization Method	Sorb/Temp	Sorbitol
Erythrocytic Oligo-level Data Points	4,406,712	1,126,762
Erythrocytic Gene-level Data Points	61,248	660,172

How to compare?

*Le Roch, KG, Y Zhou, PL Blari, M Grainger, JK Moch, JD Haynes, P De La Vega, AA Holder, S Batalov, DJ Carucci & EA Winzeler (2003) Discovery of gene function by expression profiling of the malaria parasite life cycle. Science **301**:1503-8*

*Bozdech, Z, M Llinas, BL Pulliam, ED Wong, J Zhu & JL DeRisi (2003) The transcriptome of the intraerythrocytic developmental cycle of Plasmodium falciparum. PLoS Biol. **Epub**: Aug 18.*

S ER LR ET LT ES LT
M G

Rings   Trophs   Schizonts
1 6 12 18 24 30 36 42 48
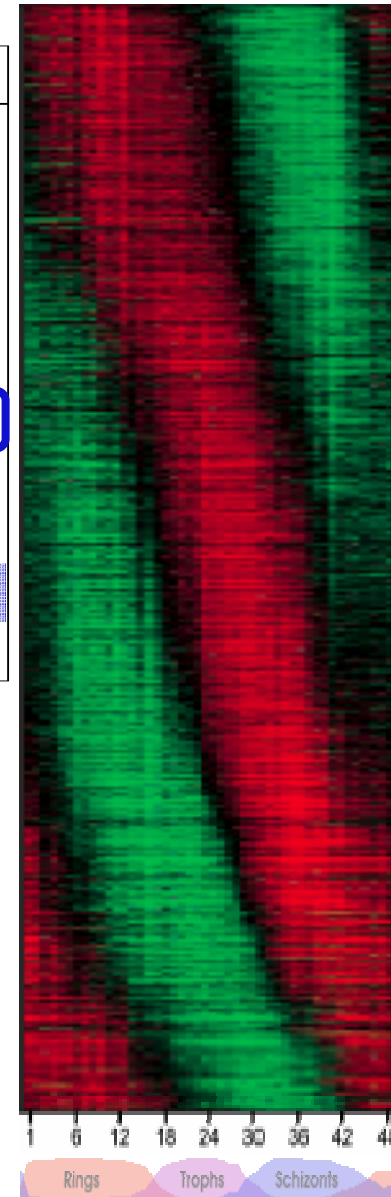
Comparing expression profiles across platforms

Le Roch *et al* **Sorb** / **Temp** data   Bozdech *et al* strain HB3 data

Absolute Expression Values
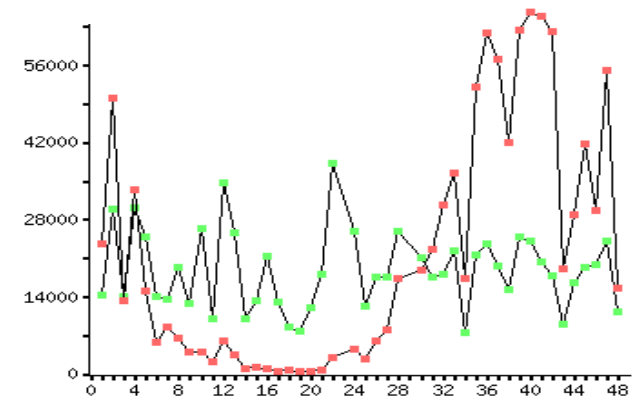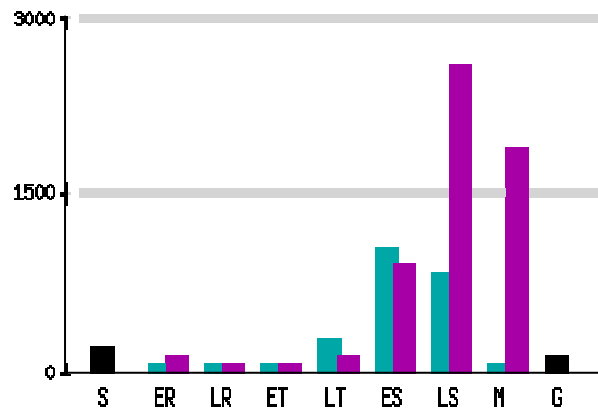
Percentiled Expression Values
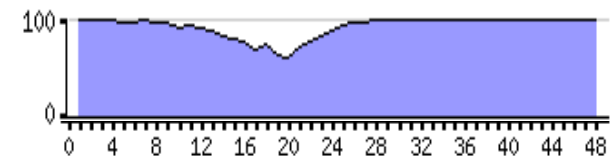
Fold Induction/Repression (log ratio expression)
GS: $\log_2$(red / green)
Affy: $\log_2(x_i / \text{ave}[x])$
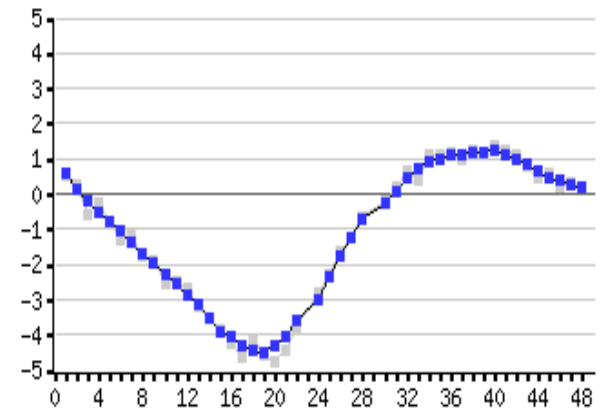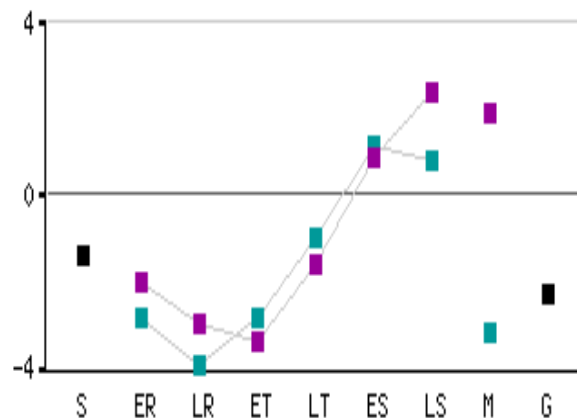
Example: MSP1

Philip Labo & Greg Grant, unpublished

Assessing profile concordance

RIFINS
(variant antigens)
n = 131

STRUCTURAL
n = 103

ALL GENES
n = 4304

random

Affy_S
GS_HB3

Fraction of Total Genes in Set

0.3

0.2

0.1

0

3000   2000   1000

low ——————— *high*

Philip Labo & Greg Grant, unpublished

PlasmoDB — The Plasmodium Genome Resource

Release 4.1
Data Release: 01/23/2003

Candidate Vaccine Antigens

Query History

Query result: rows 1 - 20

Home | Downloads | Tools | Queries | BLAST | History | CDs & Links | Browse | Data Sources | SRT

[1-20][21-26] Rows 1 - 20 of 26 [1-20][2

	gene	location	description
1	PFA0210c	pfal_chr1: 183057-184457	hypothetical protein
2	PFB0570w	pfal_chr2: 522931-523999	hypothetical protein
3	MAL6P1.299	chr6: 574310-575353	Plasmodium falciparum membrane protein pf12 precursor
4	PFE0370c	pfal_chr5: 307490-309556	subtilisin-like protease precursor, putative
5	PFE0595c	pfal_chr5: 328666-329715	hypothetical protein
6	PFE1590w	pfal_chr5: 1301219-1301764	early transcribed membrane protein
7	PFA0210c	chr1: 527658-527939	hypothetical protein
8	PF07_0128	chr7: 1265975-1270488	erythrocyte binding antigen
9	MAL6P1.299	chr9: 981838-982878	hypothetical protein
10	PFI1270w	pfal_chr9: 1040703-1041506	hypothetical protein
11	PFE0595c	pfal_chr5: 328666-329715	hypothetical protein
12	PFI1475w	pfal_chr9: 1201802-1206964	merozoite surface protein 1, precursor
13	PFI0265c	chr9: 270748-274787	hsp90 protein, putative
14	PF10_0119	chr10: 470978-471932	hypothetical protein
15	PF11_0344	chr11: 1298767-1292635	apical membrane antigen 1 precursor
16	PF10_0372	chr10: 1508412-1509473	hypothetical protein

The Genomics Unified Schema (GUS):
A Functional Genomics Data Management System

University of Pennsyl-vania

Sanger Institute

Univ. Georgia

Univ. Toronto

Java Servlets

DoTS RAD TESS SRES Core
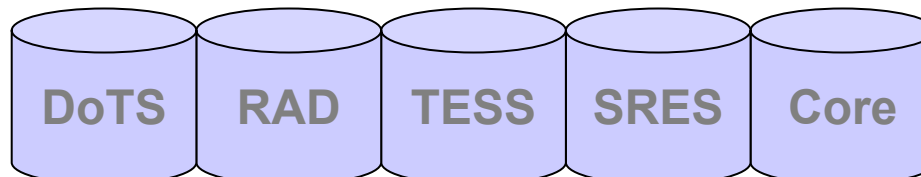
Oracle RDBMS

Object Layer for Data Loading

Flora Centromere Database

GUS

Phytophthora sojae genome

University of Chicago

Virginia Bio-informatics Institute

GUS: Genomics Unified Schema
http://www.gusdb.org

Namespace	Domain	Features
DoTS	Sequence and annotation	EST clusters Gene models
RAD	Gene Expression	MIAME/MAGE-OM
TESS	Gene Regulation	TFBS organization
Sres	Shared Resources	Ontologies
Core	Data Provenance	Documentation

GUS Components

your data
GenBank
NRDB
dbEST
SNPs
gene traps
microarrays
phenotypes
pathways
orthologs
taxonomy
GO
SO
EC
and more…

Pipeline API

Plugins (data loaders)

Data Load API

Perl Object Layer

Web Development Kit

Queries & Analysis

Warehouse
(Oracle or PostgreSQL)
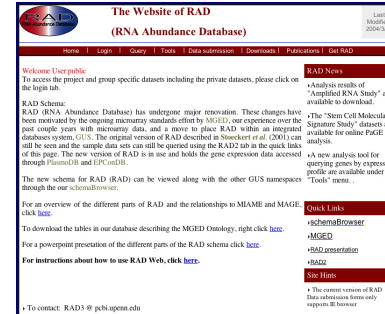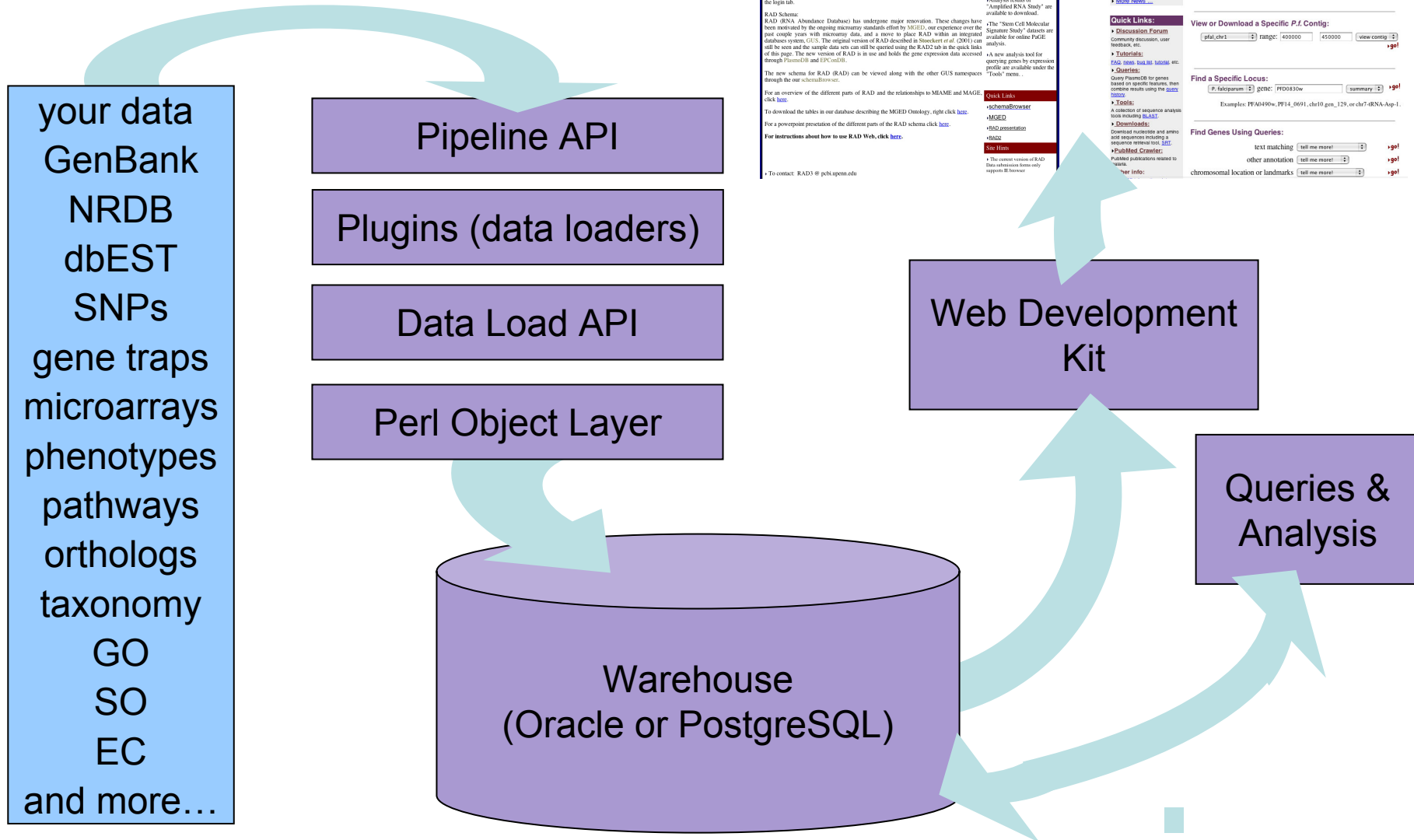
Functional Genomics with GUS

Sequence & Features

Expression (RAD)

MIAME · Study

Sample

Image Analysis

Statistical Processing

Proteomics

MIAPE · Study

Sample

Image Analysis

Statistical Processing

ImmunoHistChem

MISFISHIE · Study

Sample

Image Analysis

Statistical Processing

Central Dogma

Regulation (TESS)

Interaction

Functional Annotation of the Genome

www.mged.org psidev.sf.net www.scgap.org

Central Dogma and Sequences
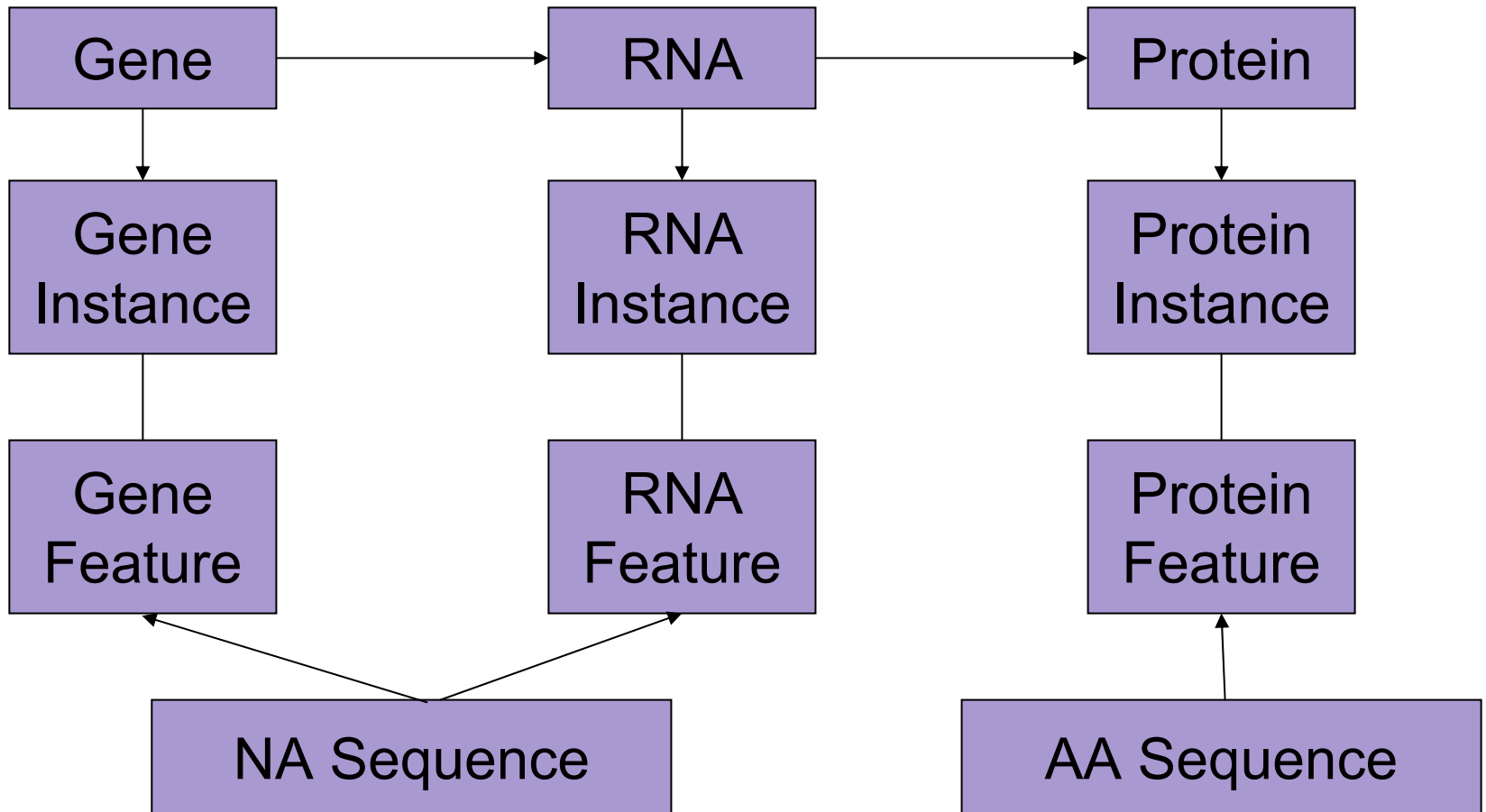
Finding all "kinases" in P. falciparum

Genes with a GO function assignment

Home   Downloads   Tools   Queries   BLAST   History   CDs & Links   Browse   Data Sources   SRT   Help

Description

This query allows one to retrieve all the genes that share the same predicted GO gene function assignment.

Query parameters

Chromosome: [all ▼]

GO Function:
```
iron-sulfur electron transfer carrier (GO:…
isocitrate dehydrogenase (GO:0004448)
isocitrate dehydrogenase (NADP+) (GO:0…
isoleucine–tRNA ligase (GO:0004822)
isomerase (GO:0016853)
kinase (GO:0016301)
kinase inhibitor (GO:0019210)
kinase regulator (GO:0019207)
kinesin motor (GO:0016326)
lactate dehydrogenase (GO:0004457)
lactoylglutathione lyase (GO:0004462)
```

Annotation type: [Pf sequencing center annotations]

Query options

Rows per page: [20 ▼]

Run query

[Submit] -- [Reset]

PlasmoDB Version: 4.2
PlasmoDB Date: 2004/07/15   Plasmodium Sequence Release Date: 2003/11

Query result: r…

Home   Downloads   Tools   Queries   BLAST   History   CD…

Query: Genes [type=Pf sequencing center annota… function 'kinase (GO:0016301)'

[1-20][21-40]   Rows 1 - 20 of 1…

	gene	location	descripti…
1	PFA0130c	pfal_chr1: 119275-121648	Serine/Thre…
2	PFA0380w	pfal_chr1: 315774-320559	serine/threo…
3	PFA0515w	pfal_chr1: 406893-413323	phosphatidy…
4	PFA0555c	pfal_chr1: 440804-442570	UMP-CMP…
5	PFB0150c	pfal_chr2: 149524-159660	protein kina…
6	PFB0520w	pfal_chr2: 469790-473491	protein kina…
7	PFB0605w	pfal_chr2: 541811-543807	Ser/Thr prot…
8	PFB0665w	pfal_chr2: 594196-599340	Ser/Thr prot…
9	PFB0815w	pfal_chr2: 720436-722660	protein kina…
10	PFC0060c	pfal_chr3: 74873-77071	Serine/threo…
11	PFC0105w	pfal_chr3: 110750-114766	serine/threo…
12	PFC0385c	pfal_chr3: 385812-391279	serine/threo…
13	PFC0420w	pfal_chr3: 419053-421354	calcium-dep…
14	PFC0485w	pfal_chr3: 485480-493186	protein kinase, putative
15	PFC0525c	pfal_chr3: 520829-523002	glycogen synthase kinase, putative
16	PFC0755c	pfal_chr3: 686827-691488	protein kinase, putative
17	PFC0945w	pfal_chr3: 891430-892781	protein kinase, putative
18	PFD0740w	pfal_chr4: 688793-692951	Plasmodium falciparum protein kinase, putative
19	PFD0755c	pfal_chr4: 696623-697312	adenylate kinase 1
20	PFD0865c	pfal_chr4: 794224-796323	cdc2-related protein kinase 1

Query History

Home   Downloads   Tools   Queries   BLAST   History   CDs & Links   Browse   Data Sources   SRT   Help

This page displays queries you have run in the current session. Use it to link back to the result sets of those queries, or to combine them into new result sets. Jump to the most recent query.

Click to Select	Query	Time	Result	Download	Size
☐	Genes [type=Pf sequencing center annotations chr=all] predicted to have GO function 'kinase (GO:0016301)'	11:16:23 AM	view	**Data Download**	104

To view results, click on "view" to the right of the desired query.

To download data, click on "data download", which will allow you to retrieve sequence data in FASTA format, or to select types of data to retrieve in a custom report

To combine data, first choose the queries of interest by clicking the check boxes at left, then combine the selected queries as follows:

[UNION]   [INTERSECT]   [TOP MINUS BOTTOM]   [BOTTOM MINUS TOP]

Results will appear as a new query.

Or [DELETE] the selected query results from your query history.

Union includes any genes identified by any selected query. **Intersect** includes only those genes identified by *every* selected query. **Top Minus Bottom** identifies genes in the first set that are not in any subsquent set checked. **Bottom Minus Top** includes genes in the last query checked that are not in previous queries selected.

PlasmoDB Version: 4.2
PlasmoDB Date: 2004/07/15   Plasmodium Sequence Release Date: 2003/11/04

To contact PlasmoDB, login to Support and feedback, or send email to help@plasmodb.org

NOTE: To download this complete list of genes as a FASTA file, click on the History link then follow the link marked "download" next to the desired query result set.

Gene Pages Integrate Sequence with Functional Gx

P. falciparum PFA0515w

Home Downloads Tools Queries BLAST History CDs & Links Browse Data Sources SRT Help

Plasmodium falciparum / CHR 1 / PFA0515w
phosphatidylinositol-4-phosphate 5-kinase, putative

Summary view   Add this gene to your History ▸

Annotation▸	Protein▸	Expression▸	Sequence▸

Annotation:
Curated Annotation▸ · UserComments▸ · GO Process▸ · GO Component▸ · GO Function▸ · EC number▸ · RefSeqs▸ · Metabolic Pathways▸ · MR4 Reagents▸ · Ortholog Group▸ · Ortholog Views▸ · Orthologs▸ · BLASTP non-Pf (graphic)▸ · BLASTP other (graphic)▸ · BLASTP NRDB▸

Protein:
PDB structures▸ · Structural Models▸ · Features (graphic)▸ · Pfam▸ · PROSITE▸ · TM domains▸ · SignalP▸ · PlasmoAP▸ · Motifs (graphic)▸ · Motifs▸ · Proteomics (graphic)▸ · Mass spec. data▸

Expression:
Microarrays▸ · Developmental series (clone array)▸ · Developmental series (Affy array)▸ · Deve slide · Prote · Mass

Sequence:
DNA (graphic)▸ · Exons▸ · SNPs▸ · mRNA/RNA sequence▸ · Protein sequence▸

Annotation

Curated Annotation
1 PUB_COMMENT
Possible phosphatidylinositol-4-phosphate 5-kinase. Simila phosphatidylinositol phosphate kinase 6 pipkinA SWALL: 6e-34, score: 370 27% id, and to Nicotiana rustica putative SWALL:Q9LLS0 (EMBL:AF157047) blast scores: E(): 1 thaliana putative phosphatidylinositol-4-phosphate 5-kinase (EMBL:AC005662) blast scores: E(): 2e-29, score: 331 28 PF01504,Phosphatidylinositol-4-phosphate 5-Kinase; HM to SM00330, Phosphatidylinositol phosphate kinases; Scan calcium-binding domain; ScanRegExp hit to PS00038, My signature;
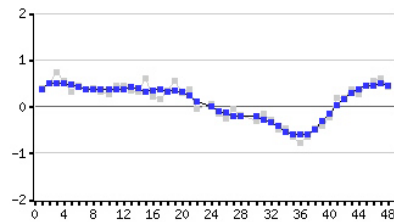
User Comments
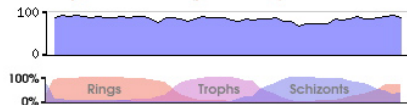0 comments available
CLICK HERE

Gene Ontology (GO) assignments

Developmental series (glass slide oligonucleotide a

Induction/Repression

PlasmoDB gene PFA0515w / Array element(s): A13725_11
DeRisi Plasmodium falciparum HB3 lifecycle Study

Expression levels (percentile)

Reference for this data: Bozdech et al. PubMed abstract
To download a free electronic reprint: go to De Risi Lab Transcriptome Database site.
See also: Supplemental material on De Risi lab. web site
See also: Study annotations in RAD

DeRisi Lab. Malaria Transcriptome Database: PFA0515w A13725_11

Query for genes whose minimum and/or maximum expression occurs at (approximately) the same time as this one:

Timing of maximal expression: [2 hours ▼] post-invasion
Maximum expression time plus or minus: [3 hours ▼]
Timing of minimum expression: [35 hours ▼] post-invasion
Minimum expression time plus or minus: [3 hours ▼]
Query parameters: [Timing of both max and min expression ▼]
Induction ratio cut-off: [>= 2-fold induction ▼]
Maximum percentile cut-off: [>= 90th percentile ▼]
Run query: Run query

The Website of RAD
(RNA Abundance Database)

Last Modified: 2004/3/11

Home | Login | Query | Tools | Data submission | Downloads | Publications | Get RAD

Welcome User: public. You are asking:
Study id: 89

There are 1 study returned

DeRisi Plasmodium Falciparum HB3 lifecycle Study (Group: CBIL, Project: PlasmodiumDB, Contact: Joe DeRisi)

Note: It may take a while to display or download the raw or normalized dataset

DeRisi Plasmodium Falciparum HB3 lifecycle Study ↑ [study info] [array info]

Annotator: Trish Whetzel
This experiment involves the analysis of the genome-wide expression profile for P. falciparum as it develops through the asexual intra-erythrocytic lifecycle in vitro. The experiment was initiated with a highly synchronous culture of parasites poised to invade new red blood cells. Upon addition of the parasites to fresh blood for invasion, we collected one hour timepoints for 48 hours to capture all of the blood-stages of the parasite. We then isolated RNA from all of these timepoints and carried out two-channel DNA microarray hybridizations using long oligonucleotide arrays. Our reference sample (Cy3 □ 532nm) is a standardized pool of RNA from most of the timepoints and represents all of the stages of development.

For an automatically generated graphic display of the biomaterials used in this study click here

• **study design** [array info]

  • Name: life-cycle time course study design with pooled reference
  • Type: time_series_design---reference_design
  • Description: One hour timepoints for 48 hours to capture all of the blood-stages of the parasite. Replicates performed for time points 1, 2, 11, 14, 18, 20, 27, 31 (duplicates except for triplicate at time point 1). Reference sample (in Cy3 channel of all arrays) is a standardized pool of RNA from most of the timepoints and represents all of the stages of development.

• **study factor**
  ○ Name: time point,
  ○ Type: ComplexAction→timepoint
  ○ Description: The length of time the samples were incubated.

  Q: Query, D: Download

The RAD Study Annotator

RAD Study-Annotator

To access the tutorials click here.

Annotation by
user: stoeckrt

for
project: Testing
group: Testing
study: Sample Study *(private)*

From Assay to Quantification
- **Assay**
  -- Hybridization Parameters
- **Image Acquisition**
  -- Image Files
  -- Acquisition Parameters
- **Image Quantification**
  -- Raw Data Files
  -- Quantification Parameters
- **Summary**

Study/Experiment*
(* MGED Core Ontology)
- **Study Design**
  -- Study Design Assay
- **Study Factor**
  -- Study Factor Value
- **Summary**

BioMaterial
- **Existing BioMaterials**
- **BioSource**
- **BioMaterial Characteristics**
- **Treatments**
  -- Sample pooling
  -- Sample splitting
  -- Labeling
  -- Other treatments
  -- Treatment parameters
- **AssayLabeledExtract**
- **Summary**

Miscellaneous
- **Contact**
- **Protocol**

Satisfies the MIAME checklist, using the MGED Ontology

Allows entering very specific experimental details

Web-based forms:
  Modular structure
  Written in PHP
  Front-end data integrity checks using JavaScript

Manages data privacy, based on project & group selections in GUS schema

Obtaining and Using GUS; Visions for the Future

- see www.gusdb.org; more info at www.gusdb.org/documentation
- active gusdev mailing list
- relatively straightforward to install, but loading data can be a struggle
  - growing number of tools available
  - addressing how to use and write tools with visits
  - improve 'install' scripts, documentation
  - postgres version
- Web Development Kit (WDK) to generate web sites on GUS
- extendable to all areas of functional genomics
  - sequence & array-based expression experiments
  - array CGH, 2D gel electrophoresis, MS, yeast-2-hybrid data, etc
  - *in situ* hybridizations, metabolites
- Interoperability with other GUS installations and common tools
  - exchange files and scripts, MAGE-ML (use community standards
  - web services (exchange objects)
  - interface with open source tools (Gbrowse, Artemis, Apollo, etc)

Plans for ApiDB

- Separate GUS installations for PlasmoDB, ToxoDB, CryptoDB, and ApiDoTS (to become ApiDB)
    - each project on its own release schedule
    - common download repository for synchronizing third-party data

- Web portal providing queries across resources
    - web services to federation of GUS installations
    - queries based on common relationships
        - GO assignments
        - Ortholog and paralog groups
        - E.C. assignments and pathways
        - etc